INFORMATION PROCESSING & MANAGEMENT

An International Journal

Editor-in-Chief: **Tefko Saracevic**
Rutgers University

Volume 36 (2000) Contents and Author Index

Pergamon

INFORMATION PROCESSING & MANAGEMENT

—An International Journal—
(Incorporating *INFORMATION TECHNOLOGY*)
Email: ipm@scils.rutgers.edu
Web site: http://www.elsevier.nl/locate/infoproman

Editor-in-Chief

Tefko Saracevic
School of Communication, Information and Library Studies
Rutgers University
4 Huntington Street
New Brunswick, NJ 08903, U.S.A.
tefko@scils.rutgers.edu

Associate Editors

Nicholas J. Belkin
School of Communication, Information and Library Studies
Rutgers University
4 Huntington Street
New Brunswick, NJ 08901-1071, U.S.A.
nick@belkin.rutgers.edu

W. Bruce Croft
Department of Computer and Information Science
University of Massachusetts
Amherst, MA 01003, U.S.A.
croft@cs.umasss.edu

Associate Editor (Europe)

C. J van Rijsbergen
Department of Computing Science
Lilybank Gardens
The University
Glasgow G12 8QQ, Scotland
keith@dcs.gla.ac.uk

Associate Editor (Book Reviews)

Amanda Spink
Associate Professor
School of Information Sciences & Technology
The Pennsylvania State University
120 S. Burrowes Street, 511 Rider Building
University Park, PA 16801-3857, U.S.A.
spink@ist.psu.edu

Founding Editor

Harold Borko
Graduate School of Education and Information Studies
University of California
102 South Hall
Los Angeles, CA 90024-1520 U.S.A.
hborko@ucla.edu

Editorial Board

Maristella Agosti
Padova University, Padova, Italy
agosti@dei.unipd.it

Tatiana Aparac
University of Zagreb, Zagreb, Croatia
taparac@mudrac.ffzg.hr

Trudi Bellardo Hahn
University of Maryland, MD U.S.A.
th90@umail.umd.edu

David C. Blair
University of Michigan
Ann Arbor, MI U.S.A.
dcblair@sils.umich.edu

Abraham Bookstein
University of Chicago, Chicago, IL U.S.A.
a-bookstein@uchicago.edu

Michael Buckland
University of California
Berkeley, CA U.S.A.
buckland@sims.berkeley.edu

Yves Chiaramella
University of Grenoble
Grenoble Cedex, France
chiara@imag.fr

Lee Feng Chien
Academia Sinica, Taipei, Taiwan
lfchien@iis.sinica.edu.tw

Raya Fidel
University of Washington
Seattle, WA U.S.A.
fidelr@u.washington.edu

Edward A. Fox
Virginia Polytechnic Institute & State University
Blacksburg, VA U.S.A.
fox@fox.cs.vt.edu

H. P. Frei
Union Bank of Switzerland
Zurich, Switzerland
frei@ubilababs.com

Jonathan Furner
University of California
Los Angeles, CA U.S.A.
jfurner@ucla.edu

Donna Harman
National Institute of Standards & Technology
Gaithersburg, MD U.S.A.
donna.harman@nist.gov

David Harper
The Robert Gordon University
Aberdeen, Scotland
djh@scms.rgu.ac.uk

William Hersh
Oregon Health Sciences University
Portland, OR U.S.A.
hersh@ohsu.edu

Peter Ingwersen
The Royal School of Library and Information Science
Copenhagen, Denmark
pi@db.dk

Tetsuya Ishikawa
University of Library and Information Science
Taukuba, Japan
ishikawa@ulis.ac.jp

Kalervo Järvelin
University of Tampere, Tampere, Finland
kalervo.jarvelin@uta.fi

Paul B. Kantor
Rutgers University
New Brunswick, NJ U.S.A.
kantor@scils.rutgers.edu

Haruo Kimoto
Nippon Telegraph and Telephone Corporation
Yokosuka, Japan
kimoto@isl.ntt.co.jp

Gehard Knorz
Fachhochschule Darmstadt
Darmstadt, Germany
knorz@www.iud.fh-darmstadt.de

Rainer Kuhlen
Universitat Konstanz, Konstanz, Germany
rainer.kuhlen@uni-konstanz.de

Elisabeth D. Liddy
Syracuse University, Syracuse, NY U.S.A.
liddy@mailbox.syr.edu

John Lindsay
Kingston University
Kingston upon Thames, U.K.
lindsay@king.ac.uk

Elisabeth Logan
Florida State University
Tallahassee, FL U.S.A.
logan@lis.fsu.edu

Gary Marchionini
University of North Carolina
Chapel Hill, NC U.S.A.
march@ils.unc.edu

Richard S. Marcus
Massachusetts Institute of Technology
Cambridge, MA U.S.A.
marcus@lids.mit.edu

Michel J. Menou
Consultant in Information Management
Les Rosiers sur Loire, France
mmenou@imaginet.fr

Jessica L. Milstead
The JELEM Company
Indian Head, MD, U.S.A.
milstead@jelem.com

Sung Hung Myaeng
Chungnam National University
Taejon, Korea
shmyaeng@cs.chungnam.ac.kr

Desai Narasimhalu
National University of Singapore
Singapore
desai@krdl.org.sg

Fausto Rabitti
Consiglio Nazionale delle Ricerche
Pisa, Italy
F. Rabitti@cnuce.cnr.it

Roy Rada
University of Maryland
Baltimore, MD U.S.A.
rada@umbc.edu

Edie M. Rasmussen
University of Pittsburgh
Pittsburgh, PA U.S.A.
erasmus@sis.pitt.edu

Stephen E. Robertson
Microsoft Research Ltd, Cambridge, U.K.
ser@microsoft.com

Nancy Roderer
National Library of Medicine
Bethesda, MD U.S.A.
nancy_roderer@nlm.nih.gov

Karen Sparck Jones
University of Cambridge, Cambridge, U.K.
sparckjones@cl.cam.ac.uk

Author Services Department: For queries relating to the general submission of articles (including electronic text and artwork) and the status of accepted manuscripts, please contact the Author Services Department: *E-mail:* authors@elsevier.co.uk; *Fax:* +44 (0) 1865 843905; *Tel:* +44 (0) 1865 843900.

Publication information: Information Processing & Management (ISSN 0306-4573). For 2000, volume 36 is scheduled for publication. Subscription prices are available upon request from the Publisher or from the Regional Sales Office nearest you or from this journal's website (http://www.elsevier.nl/locate/infoproman). Further information is available on this journal and other Elsevier Science products through Elsevier's website: (http://www.elsevier.nl). Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by standard mail (surface within Europe, air delivery outside Europe). Priority rates are available upon request. Claims for missing issues should be made within six months of the date of dispatch.

Orders, claims, and product enquiries: please contact the Customer Support Department at the Regional Sales Office nearest you:
New York: Elsevier Science, PO Box 945, New York, NY 10159-0945, USA; phone: (+1) (212) 633 3730 [toll free number for North American customers: 1-888-4ES-INFO (437-4636)]; fax: (+1) (212) 633 3680; e-mail: usinfo-f@elsevier.com
Amsterdam: Elsevier Science, PO Box 211, 1000 AE Amsterdam, The Netherlands; phone: (+31) 20 4853757; fax: (+31) 20 4853432; e-mail: nlinfo-f@elsevier.nl
Tokyo: Elsevier Science, 9-15 Higashi-Azabu 1-chome, Minato-ku, Tokyo 106-0044, Japan; phone: (+81) (3) 5561 5033; fax: (+81) (3) 5561 5047; e-mail: info@elsevier.co.jp
Singapore: Elsevier Science, No. 1 Temasek Avenue, #17-01 Millenia Tower, Singapore 039192; phone: (+65) 434 3727; fax: (+65) 337 2230; e-mail: asiainfo@elsevier.com.sg
Rio de Janeiro: Elsevier Science, Rua Sete de Setembro 111/16 Andar, 20050-002 Centro, Rio de Janeiro - RJ, Brazil; phone: (+55) (21) 509 5340; fax: (+55) (21) 507 1991; e-mail: elsevier@campus.com.br [Note (Latin America): for orders, claims and help desk information, please contact the Regional Sales Office in New York as listed above]

© 2000 Elsevier Science Ltd. All rights reserved.

Periodicals postage paid at Rahway, New Jersey. Information Processing & Management (ISSN 0306-4573) is published six issues per year, January, March, May, July, September and November by Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK. The annual subscription in the USA is $919. Information Processing & Management is circulated by Mercury International Limited, 365 Blair Road, Avenel, NJ 07001, USA.
POSTMASTER: Please send address corrections to: Information Processing & Management, c/o Customer Services, Elsevier Science Inc., 655 Avenue of the Americas, New York, NY 10010, USA.

LIST OF CONTENTS

Number 1

SPECIAL ISSUE

The Sixth Text REtrieval Conference (TREC-6)

E. M. Voorhees	1	Editorial
Ellen M. Voorhees **Donna Harman**	3	Overview of the Sixth Text REtrieval Conference (TREC-6)
Karen Sparck Jones	37	Further reflections on TREC
Alan F. Smeaton	87	TREC-6: personal highlights
S. E. Robertson **S. Walker** **M. Beaulieu**	95	Experimentation as a way of life: Okapi at TREC
Chris Buckley **Mandar Mitra** **Janet Walz** **Claire Cardie**	109	Using clustering and SuperConcepts within SMART: TREC 6
Gordon V. Cormack **Charles L. A. Clarke** **Christopher R. Palmer** **Samuel S. L. To**	133	Passage-based query refinement (MultiText experiments for TREC-6)
Jose Perez-Carballo **Tomek Strzalkowski**	155	Natural language information retrieval: progress report
Ashwin Rao **Allan Lu** **Ed Meier** **Salahuddin Ahmed** **Daniel Pliske**	179	Query processing in TREC-6
Jan O. Pedersen **Craig Silverstein** **Christopher C. Vogt**	187	Verity at TREC-6: out-of-the-box and beyond

Number 2

SPECIAL ISSUE SECTION
Web-Based Information Retrieval Research

Amanda Spink **Jian Qin**	205	Editorial
Bernard J. Jansen **Amanda Spink** **Tefko Saracevic**	207	Real life, real users, and real needs: a study and analysis of user queries on the web
Peiling Wang **William B. Hawk** **Carol Tenopir**	229	Users' interaction with World Wide Web resources: an exploratory study using a holistic approach
Bert J. Dempsey **Robert C. Vreeland** **Robert G. Sumner Jr** **Kiduk Yang**	253	Design and empirical evaluation of search software for legal professionals on the WWW
Joanne Capstick **Abdel Kader Diagne** **Gregor Erbach** **Hans Uszkoreit** **Anne Leisenberg** **Manfred Leisenberg**	275	A system for supporting cross-lingual information retrieval

GENERAL PAPERS

Charles L. A. Clarke **Gordon V. Cormack** **Elizabeth A. Tudhope**	291	Relevance ranking for one to three term queries
Inien Syu **S. D. Lang**	313	Adapting a diagnostic problem-solving model to information retrieval
		Book reviews
Jian Qin	331	*How Classifications Work: Problems and Challenges in an Electronic Age* by G. C. Bowker and S. L. Star
Charles Cole	333	*High Technology and Low-Income Communities: Prospects for the Positive Use of Advanced Information Technology*; by D. A. Schon, B. Sanyal, W. J. Mitchell
Mandar Mitra	336	*Summarizing Information*; by B. Endres-Niggemeyer
D. G. Roussinov	338	*Slaves of the machine*; by Gregory J. E. Rawlins

Number 3

Anne Le Calvé
Jacques Savoy 341 Database merging strategy based on logistic regression

Rila Mandala
Takenobu Tokunaga
Hozumi Tanaka 361 Query expansion using heterogeneous thesauri

Masami Shishibori
Masafumi Koyama
Makoto Okada
Jun-ichi Aoe 379 Two improved access methods on compact binary (CB) trees

Yu Guang
Yu Daren
Rong Yihong 401 The mathematical models of the periodical literature publishing process

J. Mostafa
W. Lam 415 Automatic classification using supervised learning in a medical document filtering application

Barbara M. Wildemuth
Charles P. Friedman
John Keyes
Stephen M. Downs 445 A longitudinal study of database-assisted problem solving

Diane H. Sonnenwald
Linda G. Pierce 461 Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control

Charles Cole
Pablo Cantero
Andras Ungar 481 The development of a diagnostic–prescriptive tool for undergraduates seeking information for a social science/humanities assignment. III. Enabling devices

Birger Hjørland 501 Library and information science: practice, theory, and philosophical basis

Number 4

Erica Cosijn
Peter Ingwersen 533 Dimensions of relevance

Jack P. C. Kleijnen
Willem Van Groenendaal 551 Measuring the quality of publications: new methodology and case study

M. I. Heywood
A. N. Zincir-Heywood
C. R. Chatwin 571 Digital library query clearing using clustering and fuzzy decision-making

Fabio Crestani
Puay Leng Lee 585 Searching the web by constrained spreading activation

Nikunj P. Dalal
Zane Quible
Katherine Wyatt 607 Cognitive design of home pages: an experimental study of comprehension on the World Wide Web

A. Salminen
V. Lyytikäinen
P. Tiitinen 623 Putting documents into their work context in document analysis

Hak-Gyoon Kim
Sung-Bae Cho 643 Structured storage and retrieval of SGML documents using Grove

Brigitte Endres-Niggemeyer 659 SimSum: an empirically founded simulation of summarizing

Raymond L. Major
Cliff T. Ragsdale 683 An aggregation approach to the classification problem using multiple prediction experts

Number 5

Ellen M. Voorhees 697 Variations in relevance judgments and the measurement of retrieval effectiveness

El-Sayed Atlam
Masao Fuketa
Kazuhiro Morita
Jun-ichi Aoe 717 Similarity measurement using term negative weight and its application to word similarity

Jorng-Tzong Horng
Ching-Chang Yeh 737 Applying genetic algorithms to query optimization in document retrieval

Morten Hertzum
Annelise Mark Pejtersen 761 The information-seeking practices of engineers: searching for documents as well as for people

Number 6

K. Sparck Jones
S. Walker
S. E. Robertson
779 A probabilistic model of information retrieval: development and comparative experiments—Part 1

K. Sparck Jones
S. Walker
S. E. Robertson
809 A probabilistic model of information retrieval: development and comparative experiments—Part 2

Marie Francine Moens
Jos Dumortier
841 Text categorization: the assignment of subject descriptors to magazine articles